Data Mining & Machine Learning

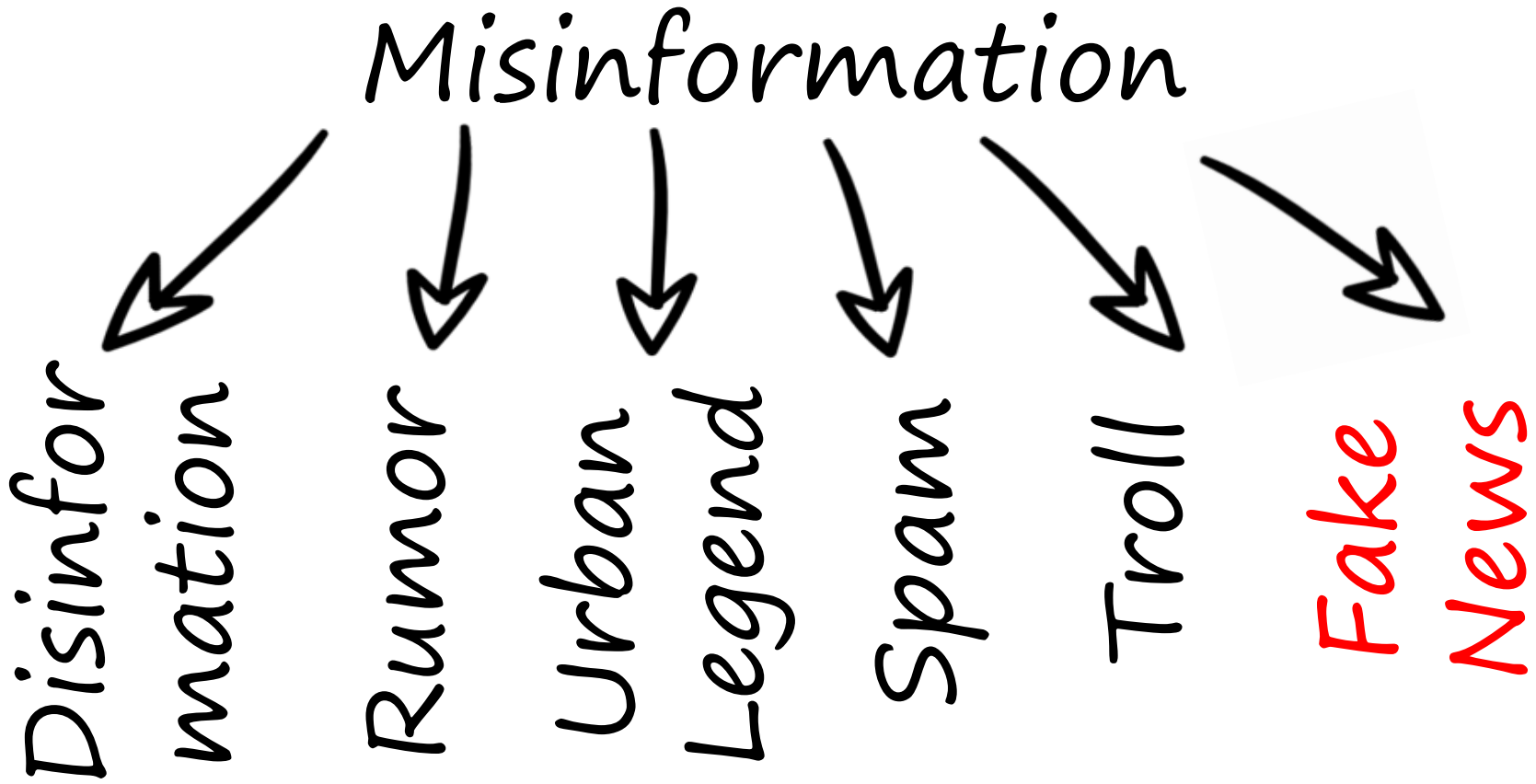
ARIZONA STATE UNIVERSITY

Social Computing Challenges in the Age of Misinformation and Fake News

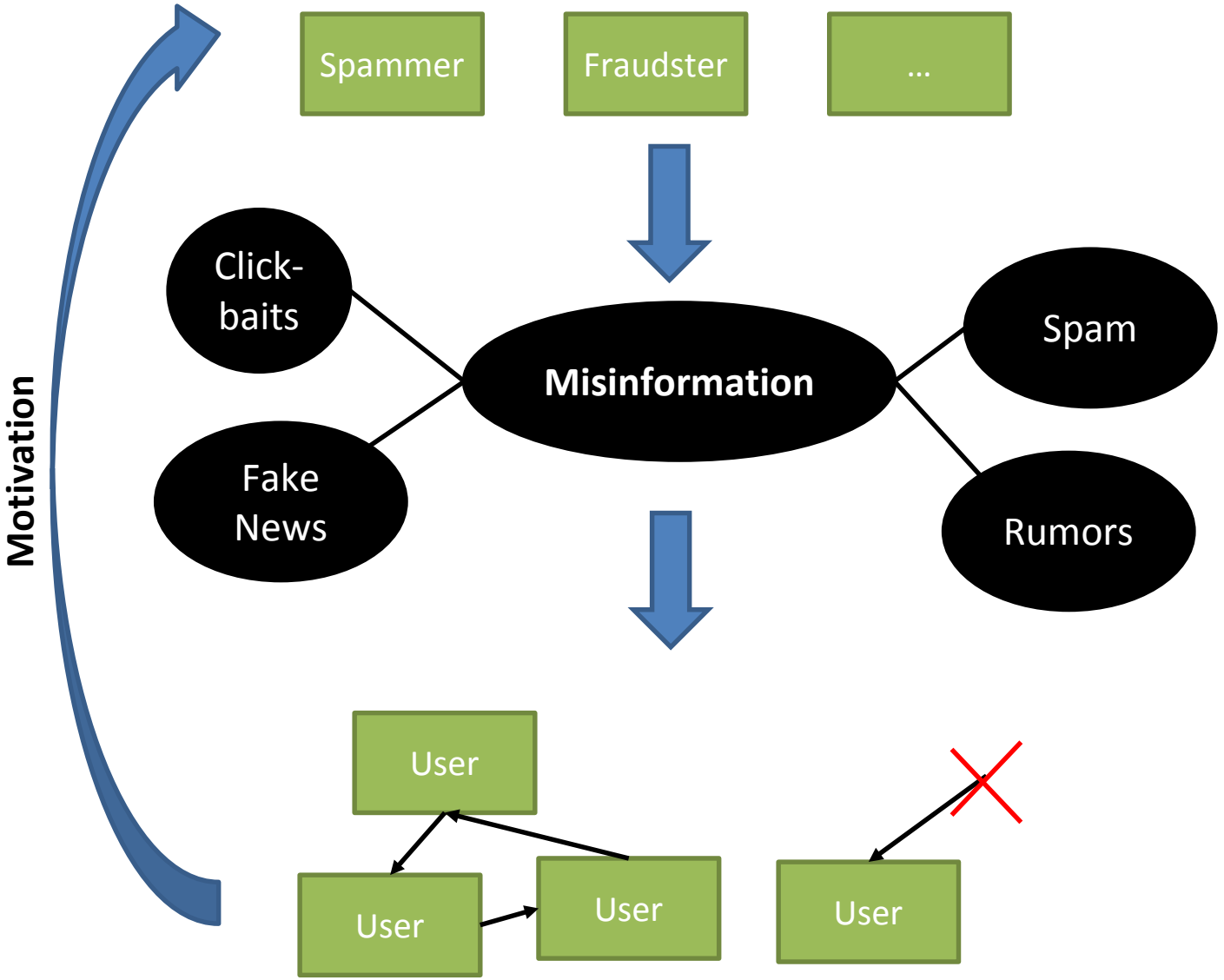
Huan Liu

Misinformation

False or inaccurate information, especially that deliberately intended to deceive.



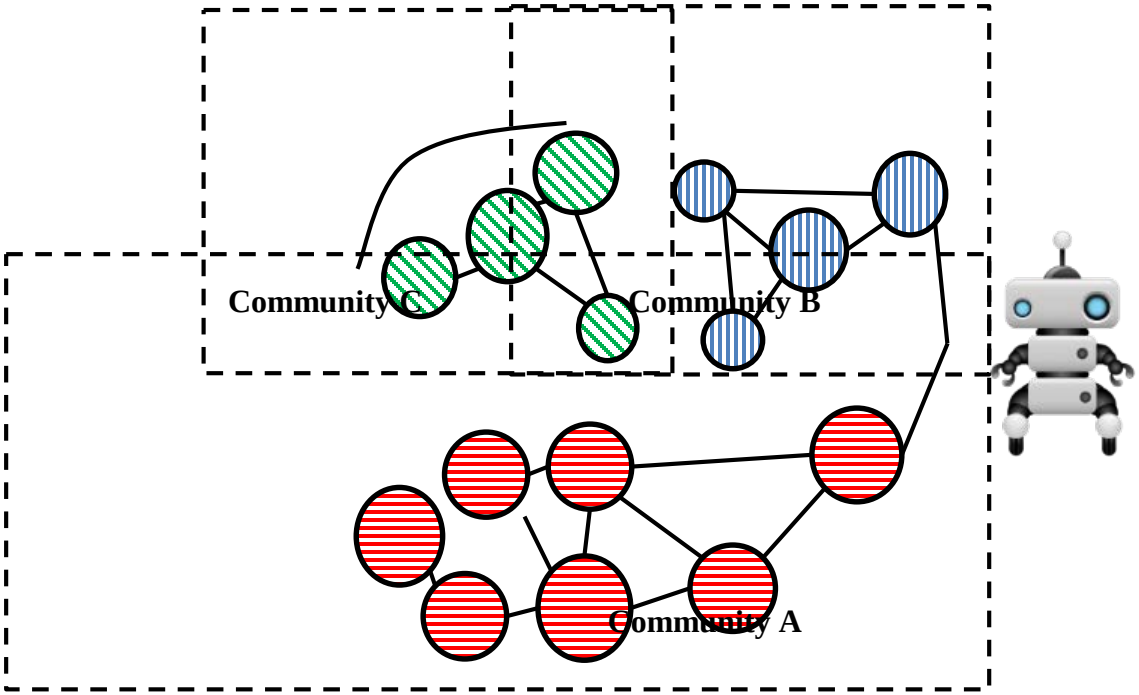
Ecosystem of Misinformation



- Spreaders
  - Spammers
  - Fraudsters
- Misinformation
  - Fake news
  - Rumor
  - ...
- Influenced Users
  - Echo chamber
  - Filter bubble

Misinformation Spreader Detection

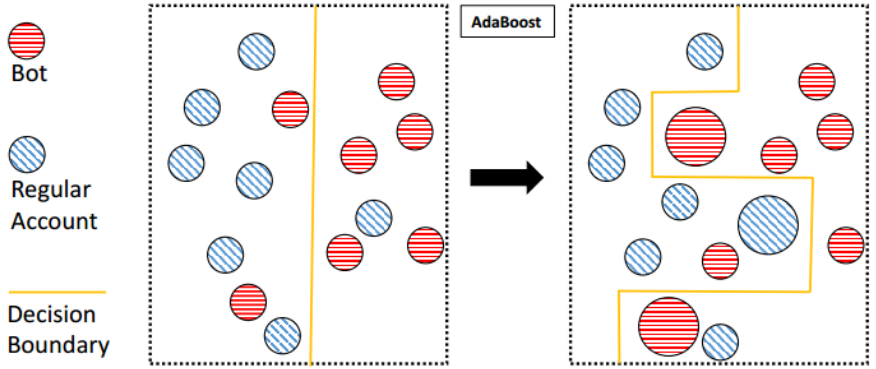
- Identifying misinformation spreaders is challenging:
  - They first disguise, and then spread misinformation
- Using community structure and group behaviors to detect flocks of misinformation spreaders



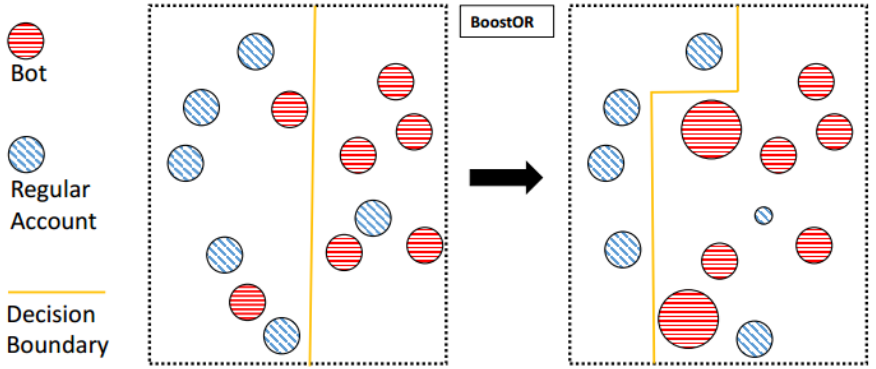
Assumption: malicious accounts cannot easily join a legitimate community

Misinformation Spreader: Bot Detection

- It is easy to create automated bot accounts
  - In our study, 9% of active Twitter users during Arab Spring are misinformation spreaders
  - 9% generates nearly 40% content
- Apply ensemble learning to improve Recall of bot detection



(a) An illustrative example of AdaBoost.



(b) An illustrative example of BoostOR.

Misinformation Detection

- Detecting by how it spreads
  - Content can be manipulated, but traces can reveal how a piece of information flows

TraceMiner: Classifying a message with its traces

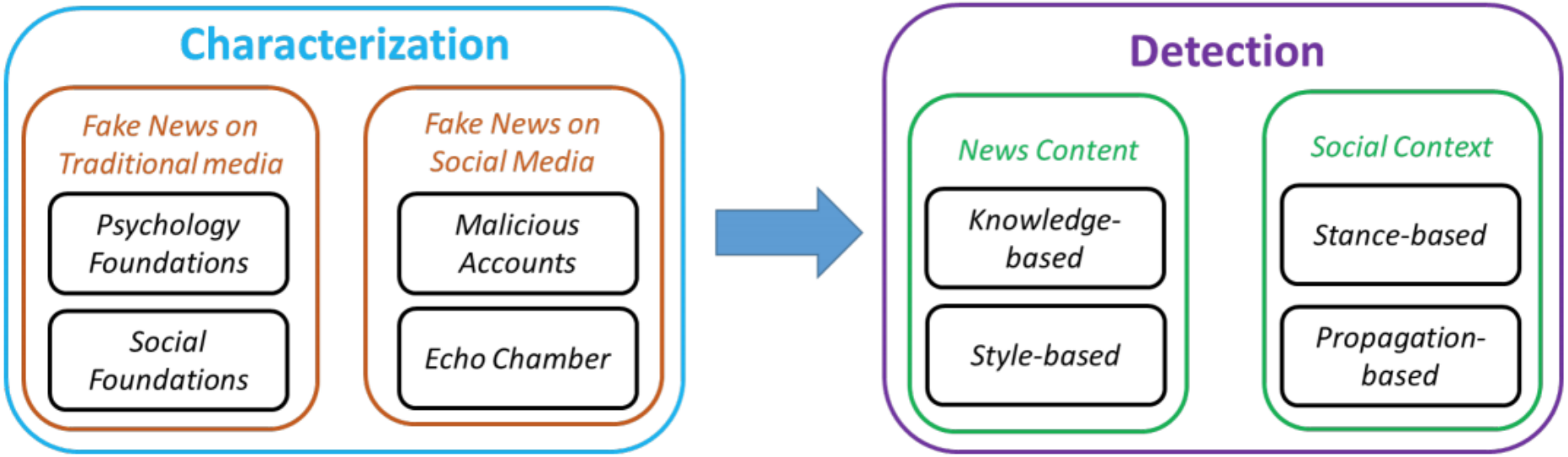
(1) Learning user embeddings with network connectivity
(2) Constructing user sequences by how message spreads
(3) Building a sequence classifier with LSTM-RNNs

Fake News

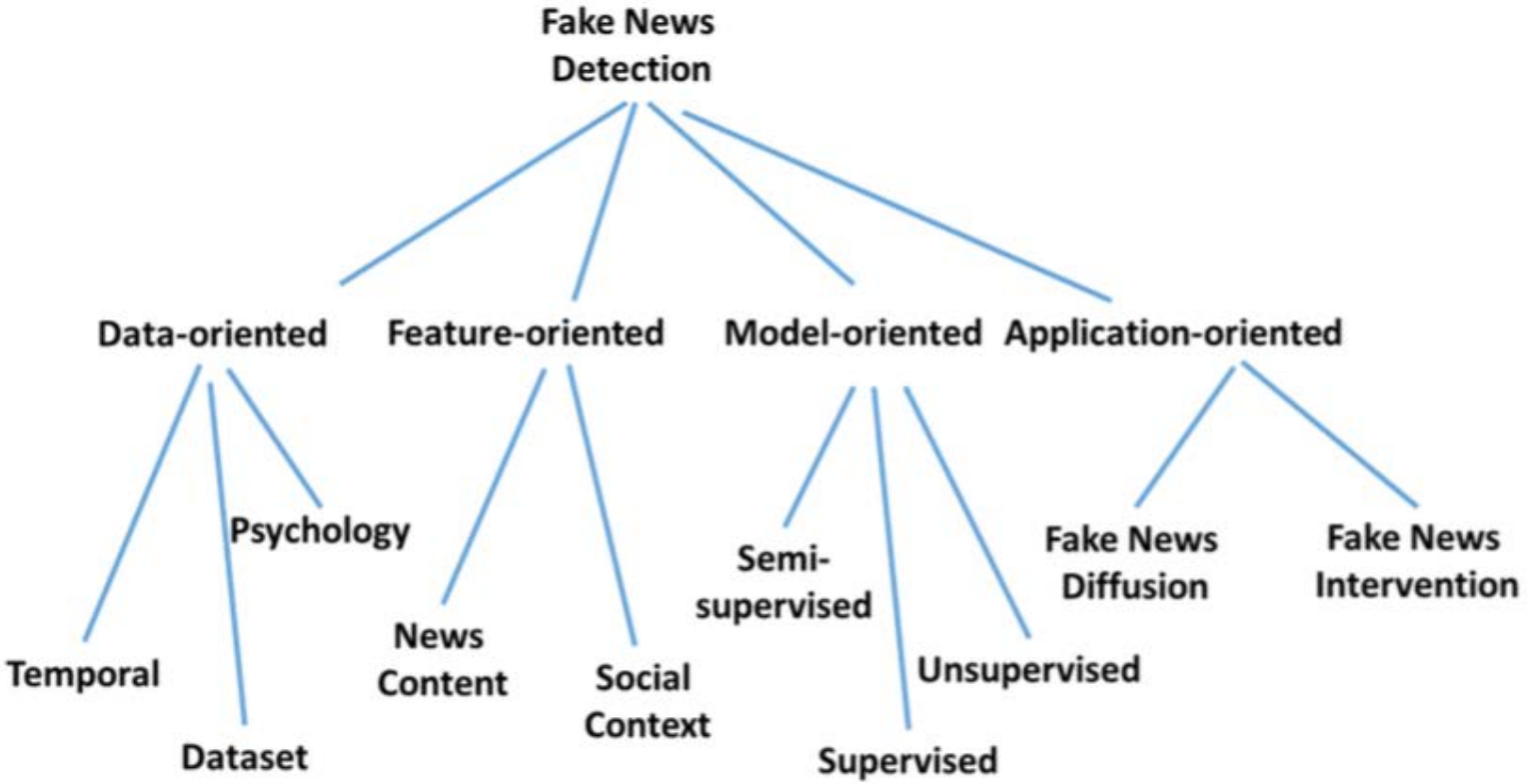
- Low quality news with intentional, false info
- People tend to seek and consume news on social media due to its low cost, easy access and rapid dissemination
- Social media accelerates the spread of fake news
- Fake news can wreak havoc
  - Mislead or confuse people
  - Change the way people consume information
    - E.g., “if I don’t like it, it is fake!”

Fake News on Social Media - Overview

- *Characterization*: Relate social and psychology theories to fake news on social media
- *Detection*: How data mining can help based on news content and social context



Research Directions of Fake News Detection

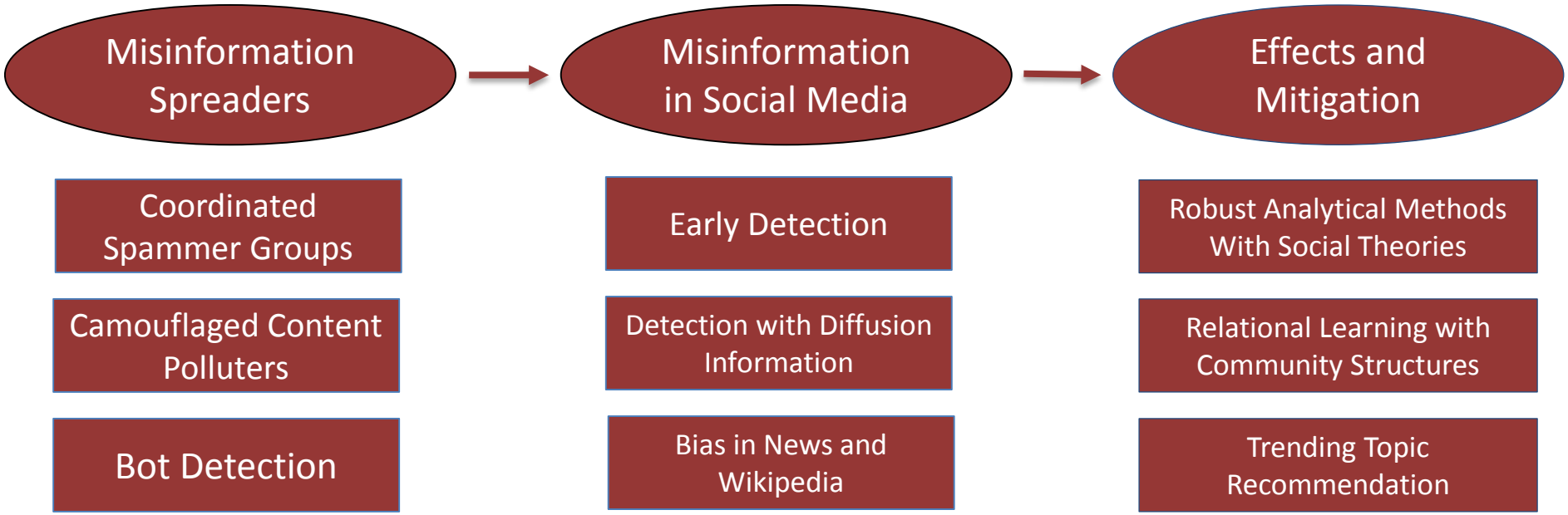


- An **Overview Paper** published in June 2017

  ``Fake News Detection on Social Media: A Data Mining Perspective", SIGKDD Explorations, 19(1)

- **Datasets**: BuzzFeedNews, LIAR, BS Detector, CREDBANDK
  - https://github.com/KaiDMML/FakeNewsNet

Mining Misinformation in Social Media



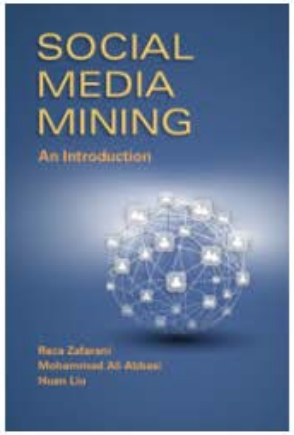
Social Media Mining

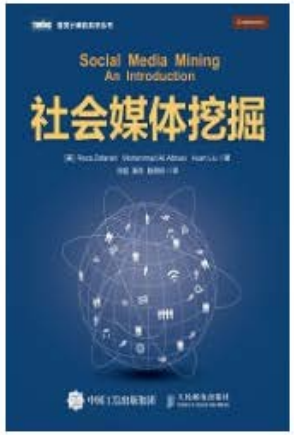
Home | Download Book | Slides/Tutorials | Table of Contents | Errata | How to Cite

Social Media Mining
An Introduction

A Textbook by Cambridge University Press

Reza Zafarani — *Syracuse University*
Mohammad Ali Abbasi — *Machine Zone*
Huan Liu — *Arizona State University*





PDF DOWNLOAD

Accessed 90,000+ times
from 160+ countries and 1200+ Universities

CAMBRIDGE UNIVERSITY PRESS | amazon.com | BARNES&NOBLE BOOKSELLERS | eBooks.com | TURING

The growth of social media over the last decade has revolutionized the way individuals interact and

http://dmml.asu.edu/smm/

THANK YOU ALL

- for this opportunity to share our research
- Acknowledgments
  - Grants from NSF, ONR, ARO, among others
  - DMML members and project leaders
  - Collaborators: CMU, CRA, Progeny, UALR

More information at
http://www.public.asu.edu/~huanliu

Further Readings

1. Liang Wu, Xia Hu, Fred Morstatter, Huan Liu, Adaptive Spammer Detection with Sparse Group Modeling. **ICWSM 2017**.
2. Liang Wu, Xia Hu, Fred Morstatter, Huan Liu. Detecting Camouflaged Content Polluters, poster paper. **ICWSM 2017**.
3. Fred Morstatter, Liang Wu, Tahora Nazer, Kathleen Carley, Huan Liu. A New Approach to Bot Detection: Striking the Balance Between Precision and Recall. **ASONAM 2016**.
4. Liang Wu, Jundong Li, Xia Hu, Huan Liu, Gleaning Wisdom from the Past: Early Detection of Emerging Rumors in Social Media. **SDM 2017**.
5. Justin Sampson, Fred Morstatter, Liang Wu, and Huan Liu, Leveraging the Implicit Structure within Social Media for Emergent Rumor Detection. **CIKM 2016**.
6. Liang Wu, Huan Liu, Tracing Fake-News Footprints: Characterizing Social Media Messages by How They Propagate. **WSDM 2018**.
7. Liang Wu, Jundong Li, Fred Morstatter, Huan Liu, Toward Relational Learning with Misinformation. **SDM 2018**.
8. Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. ``Fake News Detection on Social Media: A Data Mining Perspective", **SIGKDD Explorations**, 19(1):22-36, June, 2017.